

## Utilizarea modelor liniare unifactoriale în evaluarea costului testării software

Lect.dr. Paul POCATILU

Catedra de Informatica Economica, A.S.E. Bucuresti

*In this paper are presented some linear models for the evaluation of software testing costs. Using the software complexity, the size of the application, the productivity, the time and other factors, linear models for the evaluation of the software testing costs are built.*

**Keywords:** costs, linear evaluation models, software testing.

### Estimari si ipoteze în elaborarea modelelor liniare

Modelele liniare pentru evaluarea costurilor testării sunt elaborate pornind de la ipoteze de lucru simplificatoare, dar care nu simplifică modelele și problema. Se utilizează astfel de modele pentru că sunt ușor de construit, coeficienții se estimează cu metode pentru care există produse software și proprietățile lor permit efectuarea de analize complexe, iar cazistica verifică ipotezele specifice liniarității. Modelele își dovedesc utilitatea dacă reflectă dependente liniare între factorii exogeni și variabilele endogene din realitate pe cale experimentală.

Modelele liniare sunt *unifactoriale*, în cazul în care este o variabilă exogenă dominantă și *multifactoriale*, atunci când în model sunt luate în calcul mai multe variabile exogene.

Studierea legăturii dintre variabilă endogenă  $y$  și variabilă exogenă  $x$  este folosită la evaluarea variabilei  $y$ . Evaluarea se realizează prin ajustarea multimii de puncte printr-o linie dreaptă, aceasta fiind cunoscută sub denumirea de *dreapta de regresie*.

Pentru estimarea coeficienților modelului liniar se folosesc diferite metode, dintre care cea mai întâlnită este *metoda celor mai mici pătrate* (MCMMP). Prin MCMMP se estimează parametrii modelului liniar astfel încât suma pătratelor erorilor să fie minimă. Există și alte metode descrise în lucrări de econometrie [PECI93]:

- *metoda verosimilității maxime*; scopul metodei este de a se ajunge la valoarea cea mai probabilă pentru un parametru, rezultată din datele descrise de ecuația de regresie; metoda se utilizează în cazul în care modelele nu

sunt liniarizabile sau repartiția valorilor variabilei endogene în raport cu valorile variabilei exogene nu este normală;

- *metoda variabilelor instrumentale*; în cazul în care se constată existența de estimatori distorsionați și neconsistenți prin aplicarea MCMMP, se introduce o variabilă instrumentală cauzală în ecuație, variabilă corelată cu variabilă explicativă inițială, dar necorelată cu variabilă reziduală; parametrii noului model se determină aplicând MCMMP;

- *metoda punctelor empirice*; metoda presupune alegerea de seturi de date din seria de observații astfel încât acestea să nu fie influențate de abateri accidentale; pe baza setului de date restrâns se determină parametrii modelului;

- *metoda bayesiană*; pentru estimarea parametrilor se consideră cunoscută legea de distribuție a acestora; în cazul în care legea de distribuție a parametrilor nu este cunoscută, metoda este inaplicabilă;

- *metoda grafică*; presupune reprezentarea grafică a norului de puncte generat de perechile  $y$  și  $x$  din seria de date; prin reprezentarea grafică se aproximează parametrii și semnificativitatea acestora.

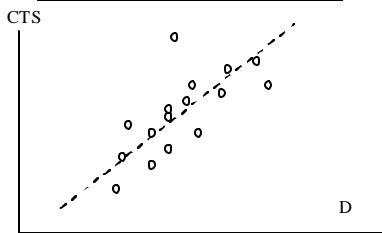
Se consideră un proces  $P$  care este influențat de factorul  $F$ . Pe baza datelor înregistrate privind evoluția în timp a variabilelor legate de procesul  $P(y)$  și factorul  $F(x)$  se construiește tabelul 1.

Pe baza înregistrărilor valorilor pentru  $Y$  și  $X$  se construiește *diagrama de împrăștiere* care arată dacă între cele două variabile există o legătură și dacă aceasta poate fi considerată liniară. Astfel, pentru studierea dependentei

liniare se realizeaza grafice corespunzatoare înregistrarilor (figurile 1 si 2).

**Tabelul 1.** Înregistrarea valorilor factorului F pentru procesul P

Y	X
y <sub>1</sub>	x <sub>11</sub>
y <sub>2</sub>	x <sub>12</sub>
...	
y <sub>t</sub>	x <sub>1t</sub>



**Fig. 1.** Corelatie pozitiva între CTS si D

$$r_{CTS/D} = \frac{\text{cov}(CTS, D)}{\sigma_D \sigma_{CTS}} = \frac{m \sum_{i=1}^m D_i CTS_i - \sum_{i=1}^m D_i \sum_{i=1}^m CTS_i}{\sqrt{\left[ m \sum_{i=1}^m D_i^2 - \left( \sum_{i=1}^m D_i \right)^2 \right] \left[ m \sum_{i=1}^m CTS_i^2 - \left( \sum_{i=1}^m CTS_i \right)^2 \right]}}$$

unde cov(CTS,D) – covarianta dintre costul testarii software(CTS) si dimensiunea aplicatiei (D);  $\sigma_{CTS}$ ,  $\sigma_D$  – abaterea medie patratica a lui CTS, respectiv D; m – numarul de înregistrari.

Coeficientul de corelatie  $r_{y/x}$  ia valori în intervalul [-1,1] si aceste valori indica daca:

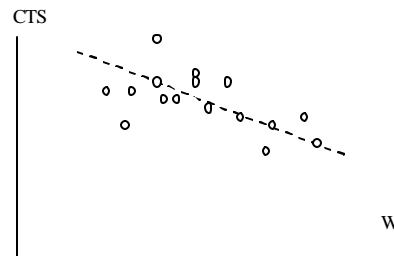
- legatura este liniara directa ( $\rightarrow 1$ );
- legatura este liniara inversa ( $\rightarrow -1$ );
- nu exista legatura ( $\rightarrow 0$ ).

Cazul  $r_{y/x} > 0$  corespunde situatiei conform careia la o crestere a nivelului variabilei x se obtine o crestere a variabilei y, iar cazul  $r_{y/x} < 0$  corespunde situatiei conform careia la o crestere a nivelului variabilei x se obtine o descrestere a nivelului variabilei y.

**Dependente liniare unifactoriale identificate în testarea software**

Se considera urmatoarii factori care influenteaza testarea software:

C	-	Complexitatea produselor software
D	-	Dimensiunea aplicatiei
NM	-	Numarul de module
GO	-	Gradul de omogenitate a echipei de realizare software
TD	-	Tehnicile de dezvoltare utilizate în realizarea produsului



**Fig. 2.** Corelatie negativa dintre CTS si W

Intensitatea legaturii dintre cele doua variabile este masurata cu ajutorul coeficientului de corelatie. Coeficientul de corelatie dintre costul testarii software si dimensiunea aplicatiei exprimata ca numar de linii sursa este dat de relatia:

SC	-	Stadiul de certificare a firmei
NU	-	Numarul utilizatorilor
PS	-	Existenta de produse similare pe piata
GN	-	Gradul de noutate software si hardware
SA	-	Specificul aplicatiei software
NI	-	Numarul elementelor de interfata
W	-	Productivitatea personalului
LP	-	Limbajul de programare utilizat
GR	-	Gradul de reutilizare software

Pornind de la forma generica a modelului liniar cu o singura variabila:  $y = ax + b$ , unde y este variabila endogena sau rezultativa si x este variabila exogena sau factoriala, se vor obtine 14 modele liniare corespunzatoare influentei unifactoriale, în care se regasesc rând pe rând factorii de influenta prezentati în lista enumerativa de mai sus.

- Model bazat pe complexitatea ciclomatica a produsului software testat:  $CTS = aC + b$ .
- Model bazat pe dimensiunea aplicatiei, exprimata în numar de linii sursa fara comentarii sau în puncte functionale:  $CTS = aD + b$ .
- Model de evaluare a costului testarii software bazat pe numarul de module sau numarul de clase ale unei aplicatii:  $CTS = aNM + b$ .

- Model liniar exprimat pe gradul de omogenitate a echipei care dezvoltă aplicația software:  $CTS = aGO + b$ .
- Model bazat pe tehnicile de dezvoltare software utilizate în realizarea produsului:  $CTS = aTD + b$ . În funcție de criteriile specifice, se acordă ponderi, ranguri sau note tehnicilor de dezvoltare software, iar valorile obținute pentru tehnicile utilizate se folosesc în modelele liniare.
- Exprimarea costului testării software în funcție de stadiul de certificare a firmei conduce la modelul liniar de evaluare:  $CTS = aSC + b$ . Stadiul de certificare a firmei se calculează în funcție de standardele aplicabile în domeniu și în funcție de nivelul de utilizare a acestora în cadrul firmei.
- Pe baza numărului utilizatorilor potențiali ai produsului software se construiește modelul liniar de evaluare a costului testării:  $CTS = aNU + b$ .
- Existența de produse similare pe piață conduce la realizarea modelului liniar unifactorial:  $CTS = aPS + b$ .
- Modelul liniar de evaluare a costului testării prin luarea în considerare a gradului de nouitate software și hardware este:  $CTS = aGN + b$ .
- Model liniar unifactorial de evaluare a costului testării software în care variabila exogenă este specificul aplicației software este:  $CTS = aSA + b$ . Tipurile de aplicații existente se cuantifică în funcție de diverse criterii, iar valorile obținute sunt utilizate în modelele de evaluare a costului testării.
- Prin luarea în calcul a numărului de elemente de interfață se obține modelul liniar de evaluare a costului testării software:  $CTS = aNI + b$ .
- Productivitatea personalului se reflectă în costul testării software. Modelul liniar de evaluare a costului testării în funcție de productivitatea personalului este:  $CTS = aW + b$ .
- Limbajul de programare utilizat stă la baza următorului model liniar de evaluare a costului testării software:  $CTS = aLP + b$ . Fiecare limbaj de programare are elemente specifice, care influențează procesul de testare. În funcție de aceste elemente, limbajele de

programare sunt cuantificate și în model sunt incluse valorile corespunzătoare limbajelor utilizate.

- Modelul de evaluare a costului testării în funcție de gradul de reutilizare software este:  $CTS = aGR + b$ . Gradul de reutilizare se exprimă ca fiind raportul dintre numărul de cazuri de test reutilizate și numărul total de cazuri de test.

În funcție de înregistrările disponibile pentru fiecare model se estimează parametrii și prin metoda celor mai mici pătrate și printr-o analiză comparată se determină validitatea și calitatea fiecăruia.

### Concluzii

Utilizarea modelelor propuse se face în funcție de datele disponibile pentru aplicația ale cărei costuri ale testării sunt evaluate. Culegerea de date este deseori dificilă, o serie de factori fiind greu de cuantificat. Modelele liniare unifactoriale de evaluare a costului testării software sunt validate și, în continuare, modelele validate sunt clasificate și împartite în grupe în funcție de criteriile de ierarhizare luate în considerare.

### Bibliografie

- [BOEH00] Boehm, Barry W. et al – *Software Cost Estimation with COCOMO II*, Prentice Hall, 2000
- [DUTA73] Duta, Lucian D., – *Manual de prezentare și utilizarea VERONICA – sistem de programe pentru calcule statistico-matematice și de prognoza*, București, 1973
- [FENT96] Fenton, Norman E., Pfleeger, Shari Lawrence – *Software Metrics, A Practical and Rigorous Approach*, International Thomson Computer Press, 1996
- [GEOR97] Georgescu, Vasile – *Econometrie*, Editura Universitaria, Craiova, 1997
- [PECI93] Pecican, Eugen Stefan – *Econometrie*, Editura All, București, 1993
- [PECI02] Pecican, Eugen, Tanasoiu, Ovidiu, Iacob Andreea – *Modele econometrice*, Biblioteca Virtuală ASE, 2002